Yuting Wei

Statistics & Data Science
307 Academic Research Building
The Wharton School, University of Pennsylvania

Email: ytwei@wharton.upenn.edu
Homepage: `https://yutingwei.github.io/`

Employment

University of Pennsylvania
- Assistant Professor, Statistics and Data Science, The Wharton School, 07/2021 - present
- Affiliated Faculty, Applied Mathematics and Computational Science (AMCS), 01/2022 - present

Carnegie Mellon University
- Assistant Professor, Statistics and Data Science, 08/2019 - 06/2021

Stanford University
- Stein Fellow/Lecturer, Department of Statistics, 09/2018 - 07/2019

Education

University of California, Berkeley, Ph.D. in Statistics, 2013 - 2018
- Thesis: A geometric perspective on some topics in statistical learning
- Advisors: Martin Wainwright, Aditya Guntuboyina
- Thesis Committee: Martin Wainwright, Aditya Guntuboyina, Peter Bickel, Venkat Anantharam
- Received *Erich L. Lehmann Citation (awarded to outstanding Ph.D. thesis in theoretical statistics)*

Peking University, Bachelor of Science (B.S.), Applied Mathematics, 2009 - 2013
Peking University, Bachelor of Arts (B.A.) Economics, 2010 - 2013

Selected Honors & Awards

Google Research Scholar Award, 2023

ICSA Junior Researcher Award, 2023

Visiting International Professor Fellowship, Ruhr-Universität Bochum, 2023

NSF CAREER Award, 2022

Bernoulli Society New Researcher Award (honorable mention), 2022

INFORMS George Nicholson Award, Finalist, 2021

Stein Fellowship, Stanford University, 2019

Erich L. Lehmann Citation, University of California, Berkeley, 2018 (awarded to an outstanding Ph.D. dissertation in theoretical statistics)

Outstanding Graduate of Peking University, 2013

Outstanding Winner of Annual Peking University Research Competition (top 10/5000), 2013

Best Overall Winner in 2nd Machine Learning Competition in Immunology, 2012

President Research Fund, Peking University, 2011-2012

Research Interests

- High-dimensional and nonparametric statistics
- Reinforcement learning theory
- Statistical learning theory
- Mathematical data science
- Applications in bioinformatics and genomics

Short-term Appointments

Simons Institute for the Theory of Computing, University of California, Berkeley

Visiting scientist, 08/2021 - 12/2021

Institute for Mathematical Research (FIM), ETH Zürich

Visiting researcher, 09/2015 - 12/2015

Department of Mathematics, City University of Hong Kong

Visiting researcher with Stephen Smale, 08/2012 - 11/2012

Journal Articles (preprints or submitted)

P1. Gen Li, **Yuting Wei**, Yuejie Chi, Yuxin Chen, "A sharp convergence theory for the probability flow ODEs of diffusion models," arXiv:2408.02320, 2024

P2. Gen Li*, Zhihan Huang*, **Yuting Wei**, "Towards a mathematical theory for consistency training in diffusion models," arXiv:2402.07802, 2024. (*= equal contributions)

P3. Yuchen Wu, Minshuo Chen, Zihao Li, Mengdi Wang, **Yuting Wei**, "Theoretical insights for diffusion guidance: A case study for Gaussian mixture models," arXiv:2403.01639, 2024.

P4. Gen Li, **Yuting Wei**, "A non-asymptotic distributional theory of approximate message passing for sparse and robust regression," arXiv:2401.03923, 2024. **(Paper awarded for 2024 ICSA Junior Researcher Award).**

P5. Gen Li, **Yuting Wei**, Yuxin Chen, Yuejie Chi, "Towards faster non-asymptotic convergence for diffusion-based generative models," arXiv:2306.09251, 2023.

P6. Gen Li, **Yuting Wei**, "A non-asymptotic framework for approximate message passing in spiked models," arxiv:2208.03313, 2023.

P7. Yue Li, **Yuting Wei**, "Minimum ℓ_1 interpolators: precise asymptotics and multiple descent," arxiv:2110.09502, 2022.

P8. Pratik Patil, Arun Kuchibhotla, **Yuting Wei**, Alessandro Rinaldo, "Mitigating multiple descents: A model-agnostic framework for risk monotonization," arxiv:2205.12937, 2022.

P9. Laixi Shi, Gen Li, **Yuting Wei**, Yuxin Chen, Matthieu Geist, Yuejie Chi, "The curious price of distributional robustness in reinforcement learning with a generative model," arXiv:2305.16589, 2023.

P10. Jingyan Wang, Ivan Stelmakh, **Yuting Wei**, Nihar B. Shah, "Debiasing evaluations that are biased by evaluations," under minor revision, *Journal of Machine Learning Research*, 2023.

Journal Articles (published or accepted)

J1. Gen Li*, Weichen Wu*, Yuejie Chi, Cong Ma, Alessandro Rinaldo, **Yuting Wei**, "High-probability sample complexities for policy evaluation with linear function approximation," *IEEE Transactions on Information Theory*, vol. 70, no. 8, pp. 5969-5999, 2024. (*= equal contributions)

J2. Gen Li, Laixi Shi, Yuxin Chen, Yuejie Chi, **Yuting Wei**, "Settling the Sample Complexity of Model-Based Offline Reinforcement Learning," *Annals of Statistics*, vol. 52, no. 1, pp. 233-260, 2024.

J3. Michael Celentano, Andrea Montanari, **Yuting Wei** (alphabetical order), "The Lasso with General Gaussian Designs with Applications to Hypothesis Testing," *Annals of Statistics*, vol. 51, no. 5, pp. 2194-2220, 2023.

J4. Gen Li, Wei Fan, **Yuting Wei**, "Approximate message passing from random initialization with applications to Z_2 synchronization," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 120, no. 31, 2023.

J5. Zhimei Ren, **Yuting Wei**, Emmanuel J. Candès, "Derandomizing Knockoffs," *Journal of the American Statistical Association (JASA)*, vol. 118, no. 542, pp. 948-958, 2023.

J6. Gen Li, **Yuting Wei**, Yuejie Chi, and Yuxin Chen, "Breaking the Sample Size Barrier in Model-Based Reinforcement Learning with a Generative Model," *Operations Research*, vol. 72, no. 1, pp. 203-221, 2024.

J7. Gen Li, Changxiao Cai, Yuxin Chen, **Yuting Wei**, Yuejie Chi, "Is Q-Learning Minimax Optimal? A Tight Sample Complexity Analysis," *Operations Research*, vol. 72, no. 1, pp. 222-236, 2024.

J8. Shicong Cen, **Yuting Wei**, Yuejie Chi, "Fast Policy Extragradient Methods for Competitive Games with Entropy Regularization," *Journal of Machine Learning Research*, vol. 25, no. 4, pp. 1-48, 2024.

J9. Gen Li, **Yuting Wei**, Yuejie Chi, Yuxin Chen, "Softmax Policy Gradient Methods Can Take Exponential Time to Converge," *Mathematical Programming*, vol. 201, pp. 707-802, 2023.

J10. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Sample Complexity of Asynchronous Q-Learning: Sharper Analysis and Variance Reduction," *IEEE Transactions on Information Theory*, vol. 68, no. 1, pp. 448-473, 2022.

J11. Minshi Peng, Yue Li, Brie Wamsley, **Yuting Wei**, Kathryn Roeder, "Integration and transfer learning of single-cell transcriptomes via cFIT," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 118, no. 10, 2021.

J12. Shicong Cen, Chen Cheng, Yuxin Chen, **Yuting Wei** and Yuejie Chi, "Fast Global Convergence of Natural Policy Gradient Methods with Entropy Regularization," *Operations Research*, vol. 70, no. 4, pp. 2563-2578, 2022. **(INFORMS George Nicholson Paper Award Finalist)**.

J13. Chen Cheng, **Yuting Wei**, and Yuxin Chen, "Tackling Small Eigen-gaps: Fine-Grained Eigenvector Estimation and Inference under Heteroscedastic Noise," *IEEE Transactions on Information Theory*, vol. 67, no. 11, pp. 7380-7419, 2021.

J14. Minshi Peng, Brie Wamsley, Andrew Elkins, Daniel M Geschwind, **Yuting Wei**, Kathryn Roeder, "Cell Type Hierarchy Reconstruction via Reconciliation of Multi-resolution Cluster Tree," *Nucleic Acids Research*, vol. 49, no. 16, 2021.

J15. **Yuting Wei** and Martin Wainwright, "The Local Geometry of Testing in Ellipses: Tight Control via Localized Kolmogorov Widths," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5110-5129, 2020.

J16. **Yuting Wei**, Billy Fang, and Martin Wainwright, "From Gauss to Kolmogorov: Localized Measures of Complexity for Ellipses," *Electronic Journal of Statistics,* vol. 14, no. 2, pp. 2988-3031, 2020.

J17. **Yuting Wei**, Martin Wainwright and Adityanand Guntuboyina, "The Geometry of Hypothesis Testing over Convex Cones: Generalized Likelihood Tests and Minimax Radii," *Annals of Statistics,* vol. 47, no. 2, pp. 994-1024, 2019.

J18. **Yuting Wei***, Fanny Yang* and Martin Wainwright, "Early Stopping for Kernel Boosting Algorithms: A General Analysis with Localized Complexities," *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6685-6703, 2019. (*= equal contributions)

J19. Tony Cai, Adityanand Guntuboyina and **Yuting Wei** (alphabetical order), "Adaptive Estimation of Planar Convex Sets," *Annals of Statistics,* vol. 46, no. 3, pp. 1018-1049, 2018.

J20. Wen-Jun Shen*, **Yuting Wei***, Xin Guo*, Stephen Smale, Hau-San Wong and Shuaicheng Li, "MHC Binding Prediction with KernelRLSpan and Its Variations," *Journal of Immunological Methods,*

vol. 406, pp. 10-20, 2014.

Refereed Conference Proceedings

C1. Kevin Tan, Wei Fan, **Yuting Wei**, "Hybrid reinforcement learning breaks sample size barriers in linear MDPs," *Neural Information Processing Systems (NeurIPS)*, 2024.

C2. Tong Yang, Shicong Cen, **Yuting Wei**, Yuxin Chen, Yuejie Chi, "Federated natural policy gradient methods for multi-task reinforcement learning," *Neural Information Processing Systems (NeurIPS)*, 2024.

C3. Yuchen Wu, Minshuo Chen, Zihao Li, Mengdi Wang, **Yuting Wei**, "Theoretical insights for diffusion guidance: A case study for Gaussian mixture models," *International Conference on Machine Learning (ICML)*, 2024.

C4. Gen Li*, Yu Huang*, Timofey Efimov, **Yuting Wei**, Yuejie Chi, Yuxin Chen, "Accelerating convergence of score-based diffusion models, provably," *International Conference on Machine Learning (ICML)*, 2024. (*= equal contributions)

C5. Gen Li, **Yuting Wei**, Yuxin Chen, Yuejie Chi, "Towards Non-Asymptotic Convergence for Diffusion-Based Generative Models," *International Conference on Learning Representations (ICLR)*, 2024.

C6. Laixi Shi, Gen Li, **Yuting Wei**, Yuxin Chen, Matthieu Geist, Yuejie Chi, "The curious price of distributional robustness in reinforcement learning with a generative model," *Neural Information Processing Systems (NeurIPS)*, 2023.

C7. Gen Li, Yuejie Chi, **Yuting Wei**, Yuxin Chen, "Minimax-optimal multi-agent RL in Markov games with a generative model," *Neural Information Processing Systems (NeurIPS)*, **oral presentation**, 2022.

C8. Laixi Shi, Gen Li, **Yuting Wei**, Yuxin Chen, Yuejie Chi, "Pessimistic Q-Learning for offline reinforcement learning: Towards optimal sample complexity," *International Conference on Machine Learning (ICML)*, 2022.

C9. Pratik Patil, **Yuting Wei**, Alessandro Rinaldo, Ryan Tibshirani, "Uniform consistency of cross-validation estimators for high-dimensional ridge regression," *International Conference on Artificial Intelligence and Statistics (AISTATS)*, **oral presentation** (10.5% of accepted papers), 2021.

C10. Shicong Cen, **Yuting Wei**, Yuejie Chi, "Fast Policy Extragradient Methods for Competitive Games with Entropy Regularization," *Neural Information Processing Systems (NeurIPS)*, 2021.

C11. Gen Li, Yuxin Chen, Yuejie Chi, Yuantao Gu, **Yuting Wei**, "Sample-Efficient Reinforcement Learning is Feasible for Linearly Realizable MDPs with Limited Revisiting," *Neural Information Processing Systems (NeurIPS)*, 2021.

C12. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Softmax Policy Gradient Methods Can Take Exponential Time to Converge," *Conference on Learning Theory (COLT)*, 2021.

C13. Jingyan Wang, Ivan Stelmakh, **Yuting Wei**, Nihar B. Shah, "Debiasing Evaluations That are Biased by Evaluations," *AAAI Conference on Artificial Intelligence*, 2021.

C14. Gen Li, Changxiao Cai, Yuxin Chen, Yuantao Gu, **Yuting Wei**, Yuejie Chi, "Tightening the Dependence on Horizon in the Sample Complexity of Q-Learning ," *International Conference on Machine Learning (ICML)*, 2021.

C15. Yue Li, Ilmun Kim, **Yuting Wei**, "Randomized Tests for High-Dimensional Regression: A More Efficient and Powerful Solution," *Neural Information Processing Systems (NeurIPS)*, 2020.

C16. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Breaking the Sample Size Barrier in Model-Based Reinforcement Learning with a Generative Model," *Neural Information Processing Systems (NeurIPS)*, 2020.

C17. Gen Li, **Yuting Wei**, Yuejie Chi, Yuantao Gu, and Yuxin Chen, "Sample Complexity of Asynchronous Q-Learning: Sharper Analysis and Variance Reduction," *Neural Information Processing Systems*

(NeurIPS), 2020.

C18. Chen Dan, **Yuting Wei** and Pradeep Ravikumar, "Sharp Statistical Guarantees for Adversarially RobustGaussian Classification," *International Conference on Machine Learning (ICML)*, 2020.

C19. **Yuting Wei***, Fanny Yang* and Martin Wainwright, "Early Stopping for Kernel Boosting Algorithms: A General Analysis with Localized Complexities," *Neural Information Processing Systems (NeurIPS)*, **spotlight presentation**, 2017. (*= equal contributions)

C20. **Yuting Wei** and Martin Wainwright, "Sharp Minimax Bounds for Testing Discrete Monotone Distributions," *International Symposium on Information Theory (ISIT)*, 2016.

Grant Support

G1. "CIF: Small: Non-Asymptotic Foundations of Approximate Message Passing," National Science Foundation (NSF), Principal Investigator, DMS-2418156, 2024 - 2027

G2. "CAREER: Statistical Learning from a Modern Perspective: Over-parameterization, Regularization, and Generalization," National Science Foundation (NSF), Principal Investigator, DMS-2143215, 2022 - 2027

G3. "Bridging Game Theory, Statistics and Optimization in Multi-Agent Reinforcement Learning," Google Research Scholar Award, Principal Investigator, 2023 - 2024

G4. "CIF: Medium: Statistical and Algorithmic Foundations of Efficient Reinforcement Learning," National Science Foundation (NSF), Principal Investigator, CIF-2106778, 2021 - 2025

G5. "Collaborative Research: Fine-Grained Statistical Inference in High Dimension: Actionable Information, Bias Reduction, and Optimality," National Science Foundation (NSF), Principal Investigator, DMS-2015447/2147546, 2020 - 2023

G6. "CIF: Small: Resource-Efficient Statistical Inference in Networked Environments," National Science Foundation (NSF), co-Principal Investigator, CCF-2007911, 2020 - 2023

Short Courses & Tutorials

S1. "Statistical and Algorithmic Foundations of Reinforcement Learning," Short Course, *Joint Conference on Statistics and Data Science*, Yunnan, China, July, 2024.

S2. "Information-Theoretic, Statistical and Algorithmic Foundations of Reinforcement Learning," Tutorial, *International Symposium on Information Theory (ISIT)*, cotaught with Y. Chi and Y. Chen, July, 2024.

S3. "Statistical and Algorithmic Foundations of Reinforcement Learning," Short Course, *Beijing International Center for Mathematical Research and School of Mathematical Science, Peking University*, 2023.

S4. "Statistical and Algorithmic Foundations of Reinforcement Learning," CE Course, *Joint Statistical Meetings (JSM) 2023*, cotaught with Y. Chi and Y. Chen, 2023.

S5. "Non-asymptotic Analysis for Reinforcement Learning," Tutorial, *ACM SIGMETRICS*, cotaught with Y. Chi and Y. Chen, 2023.

S6. "Reinforcement Learning: Fundamentals, Algorithms, and Theory," *International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2022*, cotaught with Y. Chi and Y. Chen, 2022.

S7. "Statistical and Algorithmic Foundations of Reinforcement Learning," *ICSA 2021 Applied Statistics Symposium Short Course*, cotaught with Y. Chi, Y. Chen, and Z. Zhou, 2021.

Invited Talks

T1. "Theoretical insights for diffusion guidance: A case study for gaussian mixture models," *IMS International Conference on Statistics and Data Science (ICSDS)*, Dec. 2024.

T2. "Towards a non-asymptotic understanding of diffusion-based generative models," *Workshop on Statistical Theory of Deep Neural Network Models, Brin Mathematics Research Center (Brin MRC), University of Maryland*, Nov. 2024.

T3. "Towards a non-asymptotic understanding of diffusion-based generative models," *Department Seminar, Applied Mathematics and Statistics Department, Johns Hopkins University*, Oct. 2024.

T4. "The curious price of distributional robustness in reinforcement learning with a generative model," *INFORMS Annual Meeting, Seattle*, Oct. 2024.

T5. "Towards a non-asymptotic understanding of diffusion-based generative models," *Statistics and Data Science Seminar, Auburn University*, Oct. 2024.

T6. "Approximate message passing: A non-asymptotic framework and beyond," *Mathematical Data Science Seminar, Department of Mathematics, University of Maryland*, Sep. 2024.

T7. "Theoretical insights for diffusion guidance: A case study for gaussian mixture models ," *Bernoulli-IMS 11th World Congress in Probability and Statistics, Bochum, Germany*, Aug. 2024.

T8. "Towards a non-asymptotic understanding of diffusion-based generative models," *Workshop: Diffusions in machine learning: Foundations, generative models and non-convex optimisation, The Alan Turing Institute, virtual*, Jul. 2024.

T9. "Approximate message passing: A non-asymptotic framework and beyond," *ICSA China Conference, Wuhan*, Jun. 2024.

T10. "Towards a non-asymptotic understanding of diffusion-based generative models," *Tianyuan Mathematical Center in Central China, Wuhan University*, Jun. 2024.

T11. "Towards faster non-asymptotic convergence for diffusion-based generative models," *ICSA-Canada Chapter Symposium*, Jun. 2024

T12. "Approximate message passing: A non-asymptotic framework and beyond," *Statistics Seminar, Rutgers University*, Dec. 2023.

T13. "Approximate message passing: A non-asymptotic framework and beyond," *8th Princeton Day of Statistics, Princeton University*, Nov. 2023.

T14. "Approximate message passing: A non-asymptotic framework and beyond," *Math and Data (MaD) seminar, New York University*, Oct. 2023.

T15. "On the effectiveness and ineffectiveness of policy optimization for reinforcement learning" *UPenn Optimization Seminar*, Oct, 2023.

T16. "Towards optimal sample complexities in offline reinforcement learning and Markov games," *Online Seminar Series of Clubear*, Sep. 2023.

T17. "Settling the sample complexity of model-based offline reinforcement learning," *Joint Statistical Meetings (JSM), Toronto*, Aug. 2023.

T18. "A non-asymptotic framework for the approximate message passing algorithm," *ICSA International Conference, Hong Kong*, Jul. 2023.

T19. "The distribution of Lasso and its applications: Arbitrary covariance," *ICSA China Conference, Chengdu*, Jul. 2023.

T20. "Approximate message passing: A non-asymptotic framework and beyond," *Emmanuel Candès' Group Meeting, Stanford University*, Jun. 2023.

T21. "A non-asymptotic framework for the approximate message passing algorithm with application to Z_2 synchronization," *2023 Workshop on Statistical Network Analysis and Beyond, Anchorage*, Jun. 2023.

T22. "Breaking the sample size barrier in reinforcement learning," *Computational-Statistical Interplay in Machine Learning, MIT*, May. 2023.

T23. "A non-asymptotic framework for the approximate message passing algorithm," *SIAM Conference on Optimization, Seattle*, May. 2023.

T24. "A non-asymptotic framework for the approximate message passing algorithm," *Statistics Seminar, Stanford University*, Apr. 2023.

T25. "A non-asymptotic framework for the approximate message passing algorithm," *Statistics Colloquium, University of Chicago*, Mar. 2023.

T26. "A non-asymptotic framework for the approximate message passing algorithm," *Applied Probability and Risk (APR) Seminar, Columbia University*, Mar. 2023.

T27. "A non-asymptotic framework for approximate message passing in spiked models," *Statistics Annual Winter Workshop on Modern Computational Statistics, University of Florida, Gainesville*, Jan. 2023.

T28. "A non-asymptotic framework for approximate message passing in spiked models," *Computational and Methodological Statistics, London*, Dec. 2022.

T29. "Modern perspectives in high-dimensional statistics: Two recent stories," *Statistics Seminar, Tianyuan Mathematical Center in Central China, Wuhan*, Dec. 2022.

T30. "Modern perspectives in high-dimensional statistics: Two recent stories," *Statistical challenges and new solutions for large-scale complex data, Kunming, virtual*, Nov. 2022.

T31. "On the effectiveness and ineffectiveness of policy optimization for reinforcement learning," *Seminar of Statistics and Data Science, Hong Kong University of Science and Technology*, Oct. 2022.

T32. "Settling the sample complexity of model-based offline reinforcement learning," *Informs Annual Meeting, Indianapolis*, Oct. 2022.

T33. "Settling the sample complexity of model-based offline reinforcement learning," *Asilomar Conference on Signals, Systems, and Computers*, Oct. 2022.

T34. "A non-asymptotic framework for approximate message passing in spiked models," *SIAM Conference on Mathematics of Data Science (MDS22), San Diego*, Sep. 2022.

T35. "A non-asymptotic framework for approximate message passing in spiked models," *Probability Seminar, UC Davis*, Sep. 2022.

T36. "Beyond $o(\log n/\log\log n)$ iterations: A non-asymptotic framework for approximate message passing in spiked models," *Joint Statistical Meetings*, Washington, D.C., Aug. 2022.

T37. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Bernoulli Young Researchers Workshop, virtual*, Jul. 2022.

T38. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *ICSA China Conference, virtual*, Jul. 2022.

T39. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Youth in High-Dimensions: Recent Progress in Machine Learning, High-Dimensional Statistics and Inference*, Jun. 2022.

T40. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *International Conference on Econometrics and Statistics (EcoSta)*, Jun. 2022.

T41. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *New Advances in Statistics and Data Science, Hawaii*, May. 2022.

T42. "Beyond $o(\log n/\log\log n)$ Iterations: Non-asymptotic Analysis for Approximate Message Passing," *ICMS workshop, Edinburgh*, May. 2022.

T43. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Statistics Seminar, Cambridge University*, May. 2022.

T44. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Probabilitas Seminar, Harvard University*, Mar. 2022.

T45. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *52nd Annual Conference on Information Systems and Sciences (CISS) conference, Princeton*, Mar. 2022.

T46. "Exponential lower bounds and fast convergence for policy optimization," *Martin Wainwright's group,*

UC Berkeley, Nov. 2021.

T47. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent," *Workshop on Seeking Low Dimensionality in Deep Neural Networks (SLowDNN)*, Nov. 2021.

T48. "Breaking the sample size barrier in reinforcement learning," *Stochastics and Statistics Seminar, MIT*, Oct. 2021.

T49. "Bridge over troubled water: Lower bounds and fast convergence for policy optimization," *Informs Annual Meeting*, Oct. 2021.

T50. "Minimum ℓ_1-norm interpolators: Precise asymptotics and multiple descent" *Peter Bartlett's group, UC Berkeley*, Oct. 2021.

T51. "Sample complexity of Q-learning: Sharper analysis and minimax optimality," *ICSA 2021 Symposium*, Sep. 2021.

T52. "Breaking the sample size barrier in reinforcement learning," *Statistics Seminar, Duke University*, Sep. 2021.

T53. "Breaking the sample size barrier in reinforcement learning," *Joint Statistical Meetings (JSM)*, Aug. 2021.

T54. "Softmax policy gradient methods can take exponential time to converge," *Conference on Learning Theory (COLT)*, Aug. 2021.

T55. "Statistical Inference over Convex Cones," *SIAM Conference on Optimization*, July. 2021.

T56. "Modern statistical perspectives in reinforcement learning and early stopping," *Statistics Seminar, Harvard University*, Feb. 2021.

T57. "Breaking the sample size barrier in reinforcement learning," *Statistics Seminar, Wharton Statistics Seminar, University of Pennsylvania*, Feb. 2021.

T58. "Modern Statistical Perspectives in Reinforcement Learning and Early Stopping," *Statistics Seminar, Department of Data Sciences and Operations, University of Southern California*, Feb. 2021.

T59. "Breaking the Curse of Dimensionality: From Sparse Regression to Kernel Boosting," *Statistics Seminar, Stern School of Bussiness, New York University*, Jan. 2021.

T60. "Breaking the sample size barrier in reinforcement learning," *Statistics Seminar, Yale University*, Jan. 2021.

T61. "Breaking the Sample Size Barrier in Statistical Inference and Reinforcement Learning," *Wilks Statistics Seminar, ORFE, Princeton University*, Dec. 2020.

T62. "Breaking the Sample Size Barrier in Statistical Inference and Reinforcement Learning," *Statistics Seminar, Rutgers University*, Dec. 2020.

T63. "Breaking the sample size barrier in model-based reinforcement learning with a generative model," *Informs Annual Meeting*, Nov. 2020.

T64. "Breaking the sample size barrier in model-based reinforcement learning with a generative model," *Richard Samworth's group, Cambridge University*, Nov. 2020.

T65. "Reliable hypothesis testing paradigms in high dimensions," *Columbia University, Statistics Seminar*, Oct. 2020.

T66. "Reliable hypothesis testing paradigms in high dimensions," *Young Data Science Researcher Seminar Zurich, ETH*, Oct. 2020.

T67. "Reliable hypothesis testing paradigms in high dimensions" *Michigan State University, Statistics Seminar*, Oct. 2020.

T68. "The Lasso with general Gaussian design with application to hypothesis testing," *Joint Statistical Meeting*, Aug. 2020.

T69. "Breaking the sample size barrier in model-based reinforcement learning with a generative model," *TBSI Workshop on Learning Theory*, Shenzhen, Jul. 2020.

T70. "Understanding the distribution of the Lasso and its applications," *STATML Group, Carnegie Mellon University,* Mar. 2020.

T71. "A geometric perspective on hypothesis testing," *ICSA International Conference, Hangzhou, China,* Dec. 2019.

T72. "Towards a better understanding of the regularization in kernel learning," *TBSI Workshop on Learning Theory,* Shenzhen, Dec. 2019.

T73. "Towards a better understanding of the regularization in kernel learning," *Statistics Seminar, Cambridge University,* Nov. 2018.

T74. "Towards a better understanding of the regularization in kernel learning," *Big Data and Computational Social Science Lecture, University of British Columbia,* Mar. 2019.

T75. "Early stopping for gradient type algorithms," *Computational and Methodological Statistics, University of Pisa, Italy,* Dec. 2018.

T76. "The geometry of hypothesis testing over convex cones," *52nd Annual Conference on Information Systems and Sciences (CISS) conference, Princeton,* Mar. 2018.

T77. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, Carnegie Mellon University,* 2018.

T78. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, Stanford University,* 2018.

T79. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, University of Michigan,* 2018.

T80. "Geometric analysis of hypothesis testing and early stopping for boosting," *Statistics Seminar, University of Pennsylvania,* 2018.

T81. "Shape-constrained methods: Inference, applications, and practice," *Banff International Research Station for Mathematical Innovation and Discovery (BIRS), Canada,* 2018.

T82. "Early stopping for kernel boosting algorithms," *Neural Information Processing Systems (NIPS) conference, Long Beach,* 2017.

T83. Student talk for Oberwolfach Workshop "Statistical Recovery of Discrete, Geometric and Invariant Structures", 2017.

T84. "Sharp minimax bounds for testing discrete monotone distributions," *International Symposium on Information Theory (ISIT) conference, Barcelona,* 2016.

T85. "Adaptive estimation of planar convex sets," *Berkeley Statistics Annual Research Symposium (BSTARS),* 2016.

T86. "Sharp minimax bounds for testing discrete monotone distributions." *Stanford-Berkeley Joint Colloquium,* 2016.

Other Invited Workshops

- Semester program on computational complexity of statistical inference (long-term participant). Simons Institute, Berkeley, CA, 2021.

- Workshop on statistics meets machine learning. Oberwolfach, Germany, 2020.

- Workshop on statistical recovery of discrete, geometric and invariant structures. Oberwolfach, Germany, 2017.

Ph.D. Students & Postdocs

1. Kevin Tan, Ph.D. in Statistics and Data Science, 2023 - present (co-advised with Giles Hooker)

2. Zhanran Lin, Ph.D. in Statistics and Data Science, 2023 - present (co-advised with Zhimei Ren and Yuxin Chen)

3. Zhihan Huang, Ph.D. in Statistics and Data Science, 2023 - present (co-advised with Yuxin Chen)

4. Yu Huang, Ph.D. in Statistics and Data Science, 2023 - present (co-advised with Yuxin Chen)

5. Wei Fan, Ph.D. in Statistics and Data Science, 2022 - present

6. Weichen Wu, Ph.D. in Statistics and Data Science at CMU, 2019 - 2024
   - *first job:* The Voleon Group

7. Hong Hu, Postdoctoral researcher, 2022 - 2024
   - *placement:* Assistant Professor of Statistics and Data Science & of Electrical and Systems Engineering, Washington University in St. Louis

8. Gen Li, Postdoctoral researcher, 2021 - 2023
   - *placement:* Assistant Professor of Statistics, Chinese University of Hong Kong

9. Yue Li, Ph.D. in Statistics and Data Science at CMU, 2022
   - *first job:* Hudson River Trading

10. Minshi Peng, Ph.D. in Statistics and Data Science at CMU, 2021
    - *first job:* Amazon Inc.

Doctoral Committees

1. Pratik Patil (CMU, Statistics & Machine Learning)
2. Chen Dan (CMU, Computer Science)
3. Yufei Yi (CMU, Statistics and Data Science)

Teaching Experience

University of Pennsylvania:

- **STAT 9910-302: Mathematical Foundations of Reinforcement Learning, Fall 2023.** A new graduate course developed on statistical reinforcement learning. While reinforcement learning (RL) has garnered growing interest in recent years due to its remarkable success in practice, the statistical underpinnings of RL remain far from mature, especially when it comes to sample-starved large-dimensional regimes that are of crucial operational value in practice. This Ph.D. topic course presents a coherent framework that covers important algorithmic and statistical developments in modern RL, highlighting the connections between new ideas and classical topics.

- **STAT430: Probability, Spring 2022.**

Carnegie Mellon University:

- 36-747: Mathematics for High Dimensional Data: A statistical viewpoint, Spring 2021.
- 36-748: Mathematics for High Dimensional Data: An optimization viewpoint, Spring 2021.
- 36-225: Introduction to Probability, Fall 2020.
- 36-741: Statistics meets Optimization: Iterative sketching methods, Fall 2019.
- 36-742: Statistics meets Optimization: Approximate message passing algorithms, Fall 2019.

Stanford University:

- Statistics 314: Advanced Statistical Theory, Spring 2019.
- Statistics 206: Applied Multivariate Statistical Analysis, Winter 2018.

University and Department Service

- Hiring committee at UPenn statistics, 2023 - 2024
- Diversity search advisor at the UPenn statistics, 2023 - 2024
- Statistics seminar organizer at UPenn statistics, Fall 2023
- Ph.D. Admission committee at UPenn statistics, 2021 - 2023
- Statistics departmental postdoc selection committee at UPenn statistics, 2021 - 2023
- Statistics department seminar co-organizer at CMU, 2019 - 2021
- Faculty senate member at CMU, 2019 - 2021
- Statistics department seminar co-organizer at Stanford, 2018 - 2019

Selected Professional Service

Grant reviewer: National Science Foundation DMS Statistics program panel

Reviewer for journals: Annals of Statistics, Journal of the Royal Statistical Society: Series B, Journal of the American Statistical Association, Journal of Machine Learning Research, Biometrika, IEEE Transactions on Information Theory, Statistica Sinica, Electronic Journal of Statistics, Operations Research, SIAM Journal on Mathematics of Data Science, Journal of the Korean Statistical Society

Reviewer for conferences: Conference on Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), Annual Conference on Learning Theory (COLT), International Conference on Artificial Intelligence and Statistics (AISTATS), IEEE International Symposium on Information Theory (ISIT), Asian Conference on Machine Learning (ACML), Conference on Artificial Intelligence (AAAI)

Selected Conferences and Workshop Organization (Major Roles)

- Co-organizer of Workshop on Computational-Statistical Interplay in Machine Learning, MIT, May 2023.
- Co-organizer of the Oberwolfach Mini-workshop: Mathematical Foundations of Robust and Generalizable Learning, Germany, Oct 2022.
- Co-organizer for the AI panel of 2021 U8 World Innovation Summit, Sep 2021.